

Abstract Title Page
Not included in page count.

Title: Examining Positive but Nondifferential Gains in Secondary Students' Reading Comprehension: A Focus on Instructional Practices and Differential Benefit

Authors and Affiliations:

Primary Presenter: Deborah Simmons

Contributing Authors: Melissa Fogarty, Leslie Simmons, John Davis, & Leah Anderson
Texas A&M University

Eric Oslund
University of Texas, Austin

Email addresses of Authors

Deborah Simmons, dsimmons@tamu.edu
Melissa Fogarty, melissafogarty@tamu.edu
Eric Oslund, ericoslund@yahoo.com
Leslie Simmons, lsimmons@tamu.edu
John Davis, lokettdavis@gmail.com
Leah Anderson, leah.anderson2@gmail.com

Abstract Body

Background / Context: Adolescents in the United States and their educators face an enormous challenge with respect to reading comprehension. College and career readiness standards outlined in the Common Core State Standards Initiative (2012) place increased emphasis on preparing students to read complex text across a range of content areas. At issue is how to develop the necessary skills to be able to read the texts required of college classes and literacy-demanding occupations when fewer than 35% of students in the secondary grades read proficiently (U.S. Department of Education, 2011). Compounding the challenge is the reality that secondary-level courses are largely focused on disciplinary content where reading skills are prerequisites and therefore, not an instructional priority (Kamil et al., 2008; Kennedy & Ihle, 2012). The increased demands to read and learn from complex text in combination with the reading proficiency levels of today's adolescents bring into sharp relief the academic chasm that exists in secondary classes (Eason, Goldberg, Young, Geist, & Cutting, 2012) and the need to engage content-area teachers in finding a solution to closing it.

An emerging body of research has identified promising practices to promote reading comprehension among adolescents (Biancarosa & Snow, 2006; Boardman et al., 2008; Edmonds et al., 2009). Dole, Nokes, and Drits (2009) categorized comprehension practices into two categories: teaching practices and self-regulated learning strategies (i.e., strategies students use to gain understanding or knowledge). Findings from meta-analyses (Davis, 2010; Edmonds, 2009) suggest that multicomponent comprehension intervention holds promise. Nevertheless, to date, only a limited number of studies have examined the combination of teacher-directed and student-regulated practices conducted in secondary content-area classrooms.

Purpose / Objective / Research Question / Focus of Study: This study expands this evidence base by comparing a theoretically developed set of practices designed to develop knowledge and student-regulated text processing. Extending findings from recent meta-analyses of comprehension strategy instruction (Davis, 2010; Edmonds et al., 2009), we investigated effects of integrating teacher-directed knowledge-building and student-regulated comprehension practices on 7-10th grade students' reading comprehension. We targeted three primary research questions: (a) What are the statistical and practical effects of knowledge-building and student-regulated practices on 7th to 10th-grade students' reading comprehension? (b) What is the relationship between dimensions of fidelity of implementation (i.e., adherence, quality of instruction, and program differentiation) and reading comprehension? (c) Do knowledge-building and student-regulated practices differentially benefit subgroups of students with lower entry-level comprehension performance?

Setting: The study took place in three middle schools and three high schools from three districts in south/central Texas. School enrollments ranged from 353–1,219. All participating schools received Title I funding and had high percentages (59%-76%) of students who qualified for free or reduced-price lunch. The districts educated a diverse group of students, with 76.6% economically disadvantaged and 6.5% Limited English Proficiency (LEP) in District A, 76.2% economically disadvantaged and 15.5% LEP in District B, and 71.8% economically disadvantaged and 17.6% LEP in District C.

Population / Participants / Subjects: Seventeen 7th- to 10th-grade English language arts teachers (3 males and 14 females) participated in this study. All teachers instructed at least two sections of English language arts/reading daily ($M = 3.82$, $Mdn = 4$). Teaching experience ranged from 0–23 years with a mean of 5.66 years ($SD = 6.18$). All teachers held bachelor's degrees, three had master's degrees, and three had educational specialist or doctoral degrees. Of the 17 teachers, 14 held English language arts certifications and 7 held multiple certifications.

Students ($N = 921$) from participating classes were randomly assigned within teacher to intervention ($n = 489$; 36 classes) or typical practice conditions ($n = 432$; 29 classes). Student demographics for each condition are described in Table 1 (please insert Table 1 here). Group equivalence was evaluated using independent-sample t-tests for continuous variables (e.g., age) and chi-square tests for categorical variables (e.g., gender, ethnicity).

Intervention / Program / Practice: The intervention involved a routine of teacher-directed and student-regulated practices. To increase curricular fit, we aligned practices with the content and sequence in which text genres were introduced in each school. Teachers identified short stories and expository text as appropriate targets for comprehension intervention and recommended the intervention be restricted to the fall semester to avoid spring state assessment conflicts. Teachers identified exemplar texts for each grade (e.g. *The Most Dangerous Game* by Richard Connell and *The Monkey's Paw* by W.W. Jacobs). Organized in three phases (see Table 2), practices were introduced using short stories. The routine and instructional practices were then transferred to expository text, and finally to teacher-selected texts. Phases 1 and 2 provided structured activities and recommended language; Phase 3 was designed for generalization. Practices were implemented in the fall semester. Teachers were asked to implement practices three days per week for 50 minutes; the teachers determined their schedule of implementation within the semester.

The three primary components comprised teacher-directed Text Set-up, partner-regulated Text Analysis and Dialogue, and teacher- and student-directed Text Synthesis. Each component was used with a text (e.g., short story, expository selection), and the amount of time per component varied by text length and complexity. Narrative and informational components used a common set of practices; however, question types, graphic organizers, and assignments differed to reflect text structure and topic. The entire cycle was designed to take 120-150 minutes, distributed across three days (please insert Table 2 here)

Professional development sessions were distributed over the fall semester and included face-to-face researcher-directed sessions, small-group meetings and individual coaching. Specifically, prior to the start of the school year, research team members conducted a full day of professional development that focused on applying intervention practices to short stories. Two research staff worked with teachers over the course of the semester, meeting with them individually and in grade-level teams.

Research Design: A within-teacher, randomized block design was used to study the effects of intervention on 7th- to 10th-grade students' distal and proximal reading comprehension. Each teacher's English language arts classes were randomly assigned to either intervention or typical practice. We examined the impact of fidelity of implementation variables, including adherence,

global ratings of quality and classroom management, and program differentiation. Finally, we were interested in determining whether there was differential impact of the intervention by grade and students' entry-level comprehension.

Data Collection and Analysis: Distal Measure of Reading Comprehension (pre/posttest).

Gates MacGinitie Reading Test, 4th ed. (GMRT-4; MacGinitie, MacGinitie, Maria, Dreyer, & Hughes, 2000). The Gates MacGinitie Reading Test is a timed, group-administered assessment of reading comprehension. It consists of informational and narrative passages ranging in length from 3 to 15 sentences. Students read each passage silently and then answered 3-6 multiple-choice questions, increasing in difficulty; there is a 35-minute time limit. Internal consistency reliability ranges from .90 to .95, and alternate-form reliability is reported as .80 to .83.

Researcher-adapted proximal comprehension measure (pre/posttest). Two passages and question sets from the Adolescent Literacy Inventory (Brozo & Afflerbach, 2011) were adapted to assess proximal comprehension. The first passage was a narrative text followed by 15 multiple-choice questions. The second passage was expository with 16 multiple-choice questions. Modifications to the existing question protocols included the addition of multiple-choice answer options to existing open-ended questions, the addition of a fourth answer choice to existing multiple-choice questions, and new multiple-choice questions. Question types for both passages included literal (i.e., information presented explicitly in the text), inferential (i.e., information not explicitly stated in the text, or requiring integration of information from the text or prior knowledge), or vocabulary-based questions. The measure was untimed and was group-administered during a 50 minute session. Students earned 1 point for each correct answer with a maximum score of 31. The reliability of the measure was .80 (Cronbach's alpha).

Fidelity of implementation. Because teachers served in both conditions, we examined multiple dimensions of fidelity to evaluate implementation of intervention procedures and identify any contamination from intervention to typical practice classrooms. Fidelity measures included intervention adherence by component (intervention group only), global ratings of instruction (i.e., quality and classroom management), and program differentiation of instructional practices (Dane & Schneider, 1998).

Analyses. We used multilevel modeling to examine statistically significant differences between the intervention and the comparison condition. Three-level models were used to control for the nested nature of our data, with 921 students nested within 65 classes, which were nested within 17 teachers. The models were fit using MLwIN (MLwIN 2.20; Rasbash, Charlton, Browne, Healy, & Cameron, 2009). Outcome measures included the GMRT-4 and the researcher-adapted narrative and expository assessments. Student pretest scores on the related measures (e.g., GMRT-4 pretest for the GMRT-4 outcome variable) were used as student-level covariates to control for entry-level scores. A dummy-coded variable representing intervention and comparison condition was entered as a classroom-level variable, with 0 representing the comparison condition and 1 representing the intervention condition. Subsequent analyses used the same outcome variables and examined differences between conditions using covariates including grade level, reading status, and fidelity. The Hedge's g effect size (2007) was calculated for all models. The Benjamini-Hochberg (Benjamini & Hochberg, 1995) correction was used to control inflated family-wise error rate due to comparisons of multiple outcomes.

Findings / Results: Intervention effects. There are no statistically significant differences between conditions on any posttest measure. Hedges' g effect sizes range from $-.01$ (GMRT-4) to $.03$ (expository passage) where positive scores indicate the intervention condition performed better than comparison. There is, however, a statistically significant effect for time on the GMRT-4, with both groups evidencing higher performance at posttest than pretest ($t = 11.91$, $p < .000$). Given that pre- and posttest norms are adjusted for time, gains in performance were considered practically important with a standardized mean difference effect size equal to $.46$. The mean performance for both groups was around the 32th percentile at pretest and the 42nd percentile at posttest, with an average mean difference in standard scores of 4.35 for the intervention group and 3.84 for the comparison group. Significant pretest to posttest growth is also evident on the narrative ($t = 2.21$, $p = .027$) and expository ($t = 5.63$, $p < .000$) proximal measures. The standardized mean difference effect sizes for time are $.09$ for the narrative passage and $.22$ for the expository passage.

Fidelity effects. Mean adherence ratings averaged across phases indicate modest implementation with an average rating of 3.03 ($SD\ 1.11$). This score indicates that half of the intervention components were implemented during sessions observed. Global implementation across phases shows a slight nonsignificant decline ($r = -.144$, $p = .281$) in implementation over the course of the intervention. On average, teachers had higher overall adherence scores earlier in the intervention when using researcher-prepared lessons for short stories ($M = 3.27$). Mean adherence for the expository text phase was 3.05 and for generalization lessons, 2.86 . Simple correlations between fidelity and student achievement outcomes do not exceed -0.176 .

Findings from observations of instructional practices (not specific to intervention) indicate that teachers in both conditions employed several similar strategies to promote reading comprehension across all three phases of the intervention. For example, activating background knowledge was observed between 45 to 77% percent of typical practice conditions, in comparison, this strategy was observed between 55 to 81% of intervention conditions. Noteworthy among these findings is that teacher-regulated knowledge-building practices (background knowledge and vocabulary) looked different in the earlier phases of the study (45% typical practice and 79 - 86% intervention) but grew more alike in the generalization phase (62 - 77% typical practice; 81 - 86% intervention).

Differential effects. Results of the main effect for entry-level reading ability are mixed. On the GRMT-4, students who entered below the 15th percentile demonstrate more improvement than students who entered above that point. The opposite is true for both the narrative and expository passages. On the expository passage, students above the 15th percentile perform statistically significantly better than students below the 15th percentile (Hedges' $g = .23$). Findings from interaction analyses suggest that higher-performing readers benefitted more than their lower-performing peers from the intervention that emphasized student-regulated practices as indicated by a positive Hedges' g for the interaction term on all measures.

Conclusions: This study was conducted to advance our understanding of instructional methods to promote secondary students' comprehension of texts read in English language arts classes. The intervention that combined teacher-directed knowledge-building and student-regulated

practices did not differentially impact 7th-10th grade students' comprehension on proximal or distal measures of reading comprehension over English language arts teachers' typical practice. Effect sizes that ranged from $-.01$ to $.03$ failed to approximate those in the $.36$ - $.89$ range documented in prior research synthesis (Davis, 2010; Edmonds, 2009). Perhaps the most promising of findings is that students in both conditions evidenced statistically significant growth from pretest to posttest on proximal and standardized distal measures of reading comprehension. Our hypothesis that students with lower-entry level comprehension would benefit more than their higher-achieving peers from the integration of teacher-directed and student-regulated practices was not confirmed. In fact, students with higher-entry level reading comprehension benefitted more on narrative and expository comprehension than peers with lower entry-level comprehension.

There are three plausible explanations for the finding of statistically significant but nondifferential growth of students in the intervention and typical practice conditions. Perhaps the most probable is that the general comprehension practices used in typical practice were as effective as those specific to the intervention. A second logical explanation for the pattern of findings resides in whether there is a treatment adherence threshold necessary to produce differential performance. Our findings validate concerns raised by prior researchers about implementation variability and fidelity of implementation at the secondary level (Klingner & Vaughn, 1996; Wharton-McDonald & Swiger, 2009). Third, differential patterns of response by students with different levels of reading proficiency may have masked important patterns. We were interested in whether interventions that placed increasing emphasis on student regulation would have differential impact based on students' entry-level reading proficiency. Results of proficiency level analyses yielded nonstatistically significant but nonetheless complex effect size findings that varied by measure. On the GMRT-4, students whose entry-level scores were below the 15th percentile grew more than students above the 15th percentile with the opposite pattern occurring on the narrative and expository proximal measures. Effects sizes from interaction analyses on proximal measures indicated that students in the treatment group with higher reading scores benefitted more than students with scores below the 15th percentile with statistically significant effects on the expository measure.

Limitations. Findings of this study must be placed in the context of several limitations. First, the study was conducted in school districts with exceptionally high percentages of students with low levels of reading proficiency. Thus, further research is needed to determine whether findings generalize to schools with higher percentages of proficient readers. Second, because teachers taught both intervention and typical practice classes, they were exposed to and to some degree transferred intervention practices to their typical instruction. It is exceedingly difficult to populate randomized controlled trials in secondary schools, and teachers serving as their own control has proven to be a viable design with limited evidence of contamination. Finally, the proportion of students with significantly low reading comprehension skills was unexpected and it would be extremely valuable to understand the nature of their reading comprehension difficulties.

Appendices

Appendix A. References

- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B*, 57, 289-300.
- Biancarosa, G., & Snow, C. E. (2006). *Reading next: A vision for action and research in middle and high school literacy: A report to the Carnegie Corporation of New York (2nd Ed.)*. Washington, DC: Alliance for Excellent Education.
- Boardman, A. G., Roberts, G., Vaughn, S., Wexler, J., Murray, C. S., & Kosanovich, M. (2008). *Effective instruction for adolescent struggling readers: A practice brief*. Portsmouth, NH: RMC Research Corporation, Center on Instruction.
- Brozo, W. G. & Afflerback, P. P. (2011). *Adolescent Literacy Inventory: Grades 6-12*. Boston, MA: Pearson.
- Common Core State Standards Initiative. (2012). Common core state standards for English language arts and literacy in history/social studies, science, and technical subjects. Retrieved from (http://www.corestandards.org/assets/CCSSI_ELA%20Standards.pdf)
- Eason, S. H., Goldberg, L. F., Young, K. M., Geist, M. C., & Cutting, L. E. (2012). Reader-text interactions: How differential text and question types influence cognitive skills needed for reading comprehension. *Journal of Educational Psychology*, 3, 515-528.
[doi:10.1037/a0027182](https://doi.org/10.1037/a0027182)
- Edmonds, M. S., Vaughn, S., Wexler, J., Reutebuch, C., Cable, A., Tackett, K. K., & Schnakenberg, J. W. (2009). A synthesis of reading interventions and effects on reading

- comprehension outcomes for older struggling readers. *Review of Educational Research*, 79, 262-300. [doi:10.3102/0034654308325998](https://doi.org/10.3102/0034654308325998)
- Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychology Review*, 18(1), 23-45. [doi:10.1037/0022-0663.99.2.311](https://doi.org/10.1037/0022-0663.99.2.311)
- Davis, D. S. (2010). *A meta-analysis of comprehension strategy instruction for upper elementary and middle school students* (Unpublished doctoral dissertation). Vanderbilt University, Nashville, TN.
- Dole, J. A., Nokes, J. D., & Drits, D. (2009). Cognitive strategy instruction. In S. E. Israel & G. G. Duffy (Eds.), *Handbook of research on reading comprehension* (pp. 347-372). New York, NY: Taylor & Francis.
- Hedges, L. V. (2007). Effect sizes in cluster-randomized designs. *Journal of Educational and Behavioral Statistics*, 32, 341-370. [doi:10.3102/1076998606298043](https://doi.org/10.3102/1076998606298043)
- Kamil, M. L., Borman, G. D., Dole, J., Kral, C. C., Salinger, T., Torgesen, J., ... Spier, E. (2008). *Improving adolescent literacy: Effective classroom and intervention practices: A practice guide* (NCEE #2008-4027). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from <http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=WWC20084027>
- Kennedy, M. J., & Ihle, F. M. (2012). The old man and the sea: Navigating the gulf between special educators and the content area classroom. *Learning Disabilities Research and Practice*, 27, 44-54. [doi:10.1111/j.1540-5826.2011.00349.x](https://doi.org/10.1111/j.1540-5826.2011.00349.x)
- Klingner, J. K., & Vaughn, S. (1996). Reciprocal teaching of reading comprehension strategies

- for students with learning disabilities who use English as a second language. *Elementary School Journal*, 96, 275–293. [doi:10.1086/461828](https://doi.org/10.1086/461828)
- MacGinitie, W. H., MacGinitie, R. K., Maria, K., Dreyer, L. G., & Hughes, K. E. (2000). *Gates-MacGinitie Reading Tests* (4th ed.). Itasca, IL: Riverside.
- Rasbash, J., Charlton, C., Browne, W. J., Healy, M. & Cameron, B. (2009) *MLwiN Version 2.2*. Centre for Multilevel Modeling, University of Bristol.
- U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics (2011). *The Nation's Report Card: Reading 2011* (NCES 2012-457). Washington, DC: Author. Retrieved from <http://nces.ed.gov/nationsreportcard/pdf/main2011/2012457.pdf>
- Wharton-McDonald, R., & Swiger, S. (2009). Developing higher order comprehension in the middle grades. In S. Israel & G. Duffy (Eds.), *Handbook of Research on Reading Comprehension*, 510-530. New York: Routledge Publishers.

Appendix B. Tables and Figures

Not included in page count.

Table 1

Descriptive Student Demographics by Condition

Variable	<i>Intervention</i> (<i>n</i> = 489)		<i>Comparison</i> (<i>n</i> = 432)	
	<i>n</i>	(%)	<i>n</i>	(%)
Gender				
Male	226	46.7	206	48.2
Female	258	53.3	221	51.8
Ethnicity				
Asian	0	0.0	1	0.2
American Indian or Alaska Native	1	0.2	2	0.5
Black or African-American	160	33.2	123	28.5
Hispanic or Latino	176	36.5	148	34.7
White	145	29.7	151	35.3
Identified for special education	17	4.6	22	6.6
Free or reduced lunch	339	70.5	299	70.5
English language learner	31	6.5	29	6.8

Table 2

Intervention Emphasis and Timeline by Phase

Lesson Numbers	Type of Lessons	Suggested Timeline	Days of Instruction
Phase 1: Short Stories			
1-10	Short story introductory lessons	First Six Weeks: Weeks 4 & 5	13 days
11-13	Teacher choice of short story	First Six Weeks: Week 6	
Phase 2: Informational Text			
14-20	Informational text lessons	Second Six Weeks: Weeks 2-3	10 days
21-23	Teacher choice of informational text		
Phase 3: Generalization of Practices			
24-35	4 3-day cycles of teacher choice of text (<i>short story, informational text, novel, or combination of these text types; preferably linking a narrative and informational text</i>)	Second Six Weeks: Weeks 4-5 & Third Six Weeks: Weeks 1-2	12 days